Electrical Power and Energy Systems 103 (2018) 644–651

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Expansion planning under uncertainty for hydrothermal systems with variable resources

Benjamin Maluenda[a], Matias Negrete-Pincetic[a,c,*], Daniel E. Olivares[a,c], Álvaro Lorca[a,b,c]

[a] OCM-Lab at Department of Electrical Engineering, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile
[b] Department of Industrial and Systems Engineering, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile
[c] UC Energy Research Center, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile

ARTICLE INFO

Keywords:
Hydrothermal power systems
Power system expansion planning
Progressive hedging algorithm
Stochastic programming
Long-term uncertainty

ABSTRACT

The significant integration of variable energy resources in power systems requires the consideration of greater operational details in capacity expansion planning processes. In hydrothermal systems, this motivates a more thorough assessment of the flexibility that hydroelectric reservoirs may provide to cope with variability. This work proposes a stochastic programming model for capacity expansion planning that considers representative days with hourly resolution and uncertainty in yearly water inflows. This allows capturing high resolution operational details, such as load and renewable profile chronologies, ramping constraints, and optimal reservoir management. In addition, long-term scenarios in the multi-year scale are included to obtain investment plans that yield reliable operations under extreme conditions, such as water inflow reduction due to climate change. The Progressive Hedging Algorithm is applied to decompose the problem on a long-term scenario basis. Computational experiments on an actual power system show that the use of representative days significantly outperforms traditional load blocks to assess the flexibility that reservoir hydroelectric plants provide to the system, enabling an economic and reliable integration of variable resources. The results also illustrate the impacts of considering extreme long-term scenarios in the obtained investment plans.

1. Introduction

The large scale integration of Variable Renewable Energy (VRE) resources poses critical challenges on power system planning. In particular, the need to maintain supply and demand balanced at all times requires developing flexible and reliable power grids. Power system expansion has historically been supported by Expansion Planning (EP) tools, which have been addressed through mathematical programming for more than half a century [19]. Such optimization models need to be adapted to the new paradigm of massive integration of variable resources in power grids by re-thinking some often used assumptions and simplifications.

One of such assumptions in planning is that system load varies in a relatively predictable and slow manner, so that generation units' ramping constraints, minimum up and down times, and startup times and costs are negligible. Time is represented in these EP models through *load blocks*, which are obtained from a discretized load curve previously arranged on a decreasing order, called a *load duration curve*—typically one for each month. Electric demand and generation are then simply balanced for each load block, independently. This procedure ignores the chronology of time series and cannot accommodate unit commitment costs and constraints. Recent research and experience in systems with high VRE penetration have shown that ignoring operational constraints usually results in suboptimal investment plans [27].

Several recent works have focused on better representation of operations in EP. A novel approach is presented by Wogrin et al. [33], who discretize time into *system states* rather than load blocks. Each system state is defined by load and renewable generation level, and operational constraints are enforced between system states with a probabilistic method. A more widely used approach is the use of a *representative year* with hourly resolution for single-period investment planning. This time structure has been applied to incorporate a Unit Commitment formulation [24] and demand response [13] endogenously into EP models. This method captures the chronology of load and renewable resource profiles, and allows modeling inter-hour constraints.

A more suitable method for multi-period investment planning is the use of *representative days* with hourly resolution for each studied year, as applied by Fripp [7] and Nelson et al. [22]. This technique allows capturing hourly, seasonal, and yearly variations in load, resource

* Corresponding author at: OCM-Lab at Department of Electrical Engineering, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile.
E-mail address: mnegrete@ing.puc.cl (M. Negrete-Pincetic).
URL: http://ocm.ing.puc.cl (M. Negrete-Pincetic).

https://doi.org/10.1016/j.ijepes.2018.06.008
Received 8 September 2017; Received in revised form 26 March 2018; Accepted 1 June 2018
Available online 26 June 2018
0142-0615/ © 2018 Elsevier Ltd. All rights reserved.

Nomenclature

Sets and indices

$\Gamma_{h,s}$	Set of inflow scenarios that follow the same trajectory as inflow scenario s up to hour h
$\mathscr{B}$	Set of buses, indexed by b
$\mathscr{C}$	Set of connections in the water network, indexed by c
$\mathscr{C}_n^{in}$	Set of connections directed into water node n
$\mathscr{C}_n^{out}$	Set of connections directed out of water node n
$\mathscr{D}$	Set of representative days, indexed by d
$\mathscr{G}$	Set of all generators, indexed by g
$\mathscr{G}^H$	Set of hydro generators
$\mathscr{G}_b$	Set of generators located in bus b
$\mathscr{H}$	Set of hours, indexed by h
$\mathscr{H}_d$	Set of hours in day d
$\mathscr{H}_p$	Set of hours in period p
$\mathscr{L}$	Set of transmission lines, indexed by ℓ
$\mathscr{L}_b^{in}$	Set of transmission lines directed into bus b
$\mathscr{L}_b^{out}$	Set of transmission lines directed out of bus b
$\mathscr{N}$	Set of nodes in the water network, indexed by n
$\mathscr{N}^{\mathscr{R}}$	Set of water nodes that are reservoirs
$\mathscr{P}$	Set of investment periods, indexed by p
$\mathscr{P}2$	Set of investment periods, indexed by p
$\mathscr{S}$	Set of inflow scenarios, indexed by s

Parameters

η_ℓ^L	Transmission loss factor of line ℓ
η_g^H	Hydraulic efficiency of hydro generator $g \in G^h$ [MW/(m^3/h)]
$\overline{B^G}_g$	Investment cap per period for generator g [MW]
$\overline{B^L}_\ell$	Investment cap per period for line ℓ [MW]
$\overline{C^G}_g$	Upper bound on capacity for generator g [MW]
$\overline{C^L}_\ell$	Upper bound on capacity for line ℓ [MW]
$\overline{V}_{n,h,s}$	Upper water volume storage limit for node n at hour h and inflow scenario s [m^3]
$\phi_{g,p}^{fuel}$	Fuel cost of generator g on period p [US$/MWh]
$\phi_{g,p}^{Gfix}$	Annual fixed Operations & Maintenance (O&M) costs of generator g on period p [US$/MW/year]
$\phi_{\ell,p}^{Lfix}$	Annual fixed O&M costs of transmission line ℓ on period p [US$/MW/year]
ϕ_g^{OM}	Variable O&M costs of generator g [US$/MWh]
π_s	Probability of inflow scenario s in any year
θ_h	Scaling factor of hour h; i.e. the number of hours in a year that are represented by hour h
$\underline{V}_{n,h,s}$	Lower water volume storage limit for node n at hour h and inflow scenario s [m^3]
$b_{g,p}^G$	Existing built capacity of generator g that will be operational in period p [MW]
$b_{\ell,p}^L$	Existing built capacity of transmission line ℓ that will be operational in period p [MW]
$c_{g,h}$	Maximum generating capacity factor for generator g in hour h as fraction of installed capacity
f_p	Factor to bring costs in period p to present value
$l_{b,h}$	Load in bus b and hour h [MW]
r_g^{up}	Upward ramp rate of generator g as fraction of installed capacity
r_g^{dn}	Downward ramp rate of generator g as fraction of installed capacity
V_n^i	Initial stored water at each reservoir $n \in \mathscr{N}^{\mathscr{R}}$ [m^3]
$w_{n,h,s}$	Natural water inflow into node n at hour h and inflow scenario s [m^3/h]
y_p	Length of period p [years].

Variables

$B_{g,p}^G$	Capacity construction decision of generator g at period p [MW]
$B_{\ell,p}^L$	Capacity construction decision of line ℓ at period p [MW]
$C_{g,p}^G$	Cumulative capacity of generator g on period p [MW]
$C_{\ell,p}^L$	Cumulative capacity of line ℓ on period p [MW]
$E_{b,h,s}$	Energy curtailment in bus b at hour h under inflow scenario s [MW]
$F_{\ell,h,s}$	Power flow through line ℓ at hour h under inflow scenario s [MW]
$P_{g,h,s}$	Dispatch level of generator g at hour h under inflow scenario s [MW]
$V_{n,h,s}$	Stored water volume in water node n at hour h under inflow scenario s [m^3]
$W_{c,h,s}$	Water flow through connection c at hour h under inflow scenario s [m^3/h].

availability, and prices, leading to a better assessment of the required flexibility to accommodate high shares of VRE. Recent work by Poncelet et al. [28] analyzes methodologies to select the representative days from each year.

Another simplification still often applied in EP models that risks yielding uneconomical or unreliable plans is to take a deterministic approach and consider a single future scenario in each optimization. The volatility of energy resources' availability and cost, technological developments, and uncertain load growth motivate the endogenous inclusion of uncertainty in capacity expansion planning. Stochastic Programming (SP) has been used in EP to minimize expected costs of investment and operations in multiple scenarios. To account for operational uncertainty, work such as that by Jin et al. [12] and Park and Baldick [25] consider multiple load and wind profile scenarios with discrete probabilities.

The use of discrete scenarios has also been extended to the investment scale. A statistical procedure for load growth and fuel price scenarios is presented by Feng and Ryan [6], and expert opinion is used by Li et al. [17] to formulate climate change scenarios for multi-period EP. Munoz et al. [21] and Hobbs et al. [11] show that SP leads not only to economic plans under long-term uncertainty, but also to more reliable and adaptable systems. However, this method requires assigning discrete probabilities to each modeled scenario, which may prove a complex challenge for long-term uncertainties. Additionally, these works do not consider operational constraints that must be modeled in a chronological time framework, so flexibility requirements from VRE integration are not completely captured.

Water reservoirs in hydrothermal systems may be used to hedge against this uncertainty in multiple scales. Nevertheless, the representation of reservoir management details in EP has not received enough attention [10], due to the complexity of including constraints that link reservoir water levels throughout the time horizon, and because of the inherent uncertainty in water inflows. In systems with high VRE penetration, it becomes necessary to additionally include operational attributes of hydroelectric units, such as their high ramping capacity, to better asses the flexibility that these units may provide.

The standard to coordinate operations in hydrothermal systems, such as Chile, Sweden, Brazil, and others, is to use the Stochastic Dual Dynamic Programming (SDDP) methodology developed by Pereira and Pinto [26] or derived formulations to consider large inflow scenario trees and manage reservoirs over time. However, this method does not lend itself nicely to modeling operations in EP, since its optimal solution depends on the topology of the grid and, thus, cannot be endogenously incorporated. Some studies have used SDDP in expansion

planning models through an iterative process [3,23,31]. Nevertheless, a global optimum is not guaranteed to be reached and time resolution is represented by load blocks, which do not capture inter-hourly constraints or load and renewable profile chronology. Additionally, the iterative process implies long setup and computation times.

The large problem sizes that result from including inflow scenario trees in EP models has led most work to adopt a deterministic approach. However, some recent research includes multiple independent inflow scenarios that span throughout the whole horizon for actual power systems [4,15]. Greater detail has been captured by Gill et al. [9], who consider a set of inflow scenarios that may unfold in any given year, obtained via a scenario reduction methodology. This last approach is suitable to represent reservoir energy management for EP in systems with large storage capacity. However, as previous works, it cannot assess the flexibility that reservoirs can provide in the hourly scale because of the use of discrete load blocks. Additionally, it considers that water storage is not allowed between years, so the capability of reservoirs to hedge against inflow uncertainty is not considered.

Every mentioned EP model in industry and literature is formulated through either the perspective of a central planner that seeks to supply future demand at minimum social cost, or with a competitive framework involving markets and multiple decision makers that seek to maximize their profits [14]. Even though the generation segment is a competitive activity in most modern power markets, tools that render centralized expansion plans are still often used by agencies and regulators to study emerging system dynamics and to formulate policies. A centralized investment plan could be considered to be a proxy for expansion under perfect competition in the power market, thus providing valuable insights to decision makers [8].

The main contribution of this paper in respect to the referenced existing literature is to propose a model for EP in hydrothermal systems that captures greater operational detail and that considers uncertainty both in yearly inflows and in investment costs and inflow trends in the long-term. In contrast with previous works in hydrothermal EP, the model captures chronology of load and renewable profiles, generator ramps, and detailed reservoir management through the use of representative days with hourly resolution. A detailed modeling of hydroelectric operations is carried out, including inflow uncertainty in the yearly scale, inter-annual water storage in reservoirs, and cascading between successive generators in hydraulic basins.

Another innovative aspect of the model is the consideration of a *nominal* long-term scenario, which is the most likely to happen, and *extreme* long-term scenarios that may have severe impacts, such as inflow reduction trends and changing investment costs. The model will produce investment plans that are optimal for the nominal long-term scenario while enforcing that the system must also be reliable, though not necessarily economic, under the extreme long-term scenarios. So, the probability of the considered extreme events does not need to be quantified. The model is solved through the Progressive Hedging Algorithm (PHA) scenario decomposition technique. Each subproblem is formulated as a Linear Program, which can be efficiently solved by commercial solvers.

The remainder of the paper is organized as follows. The structure and mathematical formulation of the model are presented in Section 2. Section 3 displays the solution methodology. Computational experiments are reported in Section 4 and conclusions are finally drawn in Section 5.

2. Model formulation

Generation and transmission investment decisions are taken in every period. Each period consists of multiple years in which operational decisions, such as power dispatch, transmission flows, among others, are obtained for a number of representative days with hourly resolution. Uncertainty in the operational scale is captured in water *inflow scenarios*. A single expected load and renewable production profile is considered for each *inflow scenario*, so the capability of water reservoirs to hedge against water inflow uncertainty may be better evaluated. Uncertainty in the investment scale is captured in *long-term scenarios*, which are divided into one *nominal scenario*, for which the system will be economically optimized, and several *extreme scenarios*, which will only be taken into account to ensure reliability. Fig. 1 illustrates this decision structure.

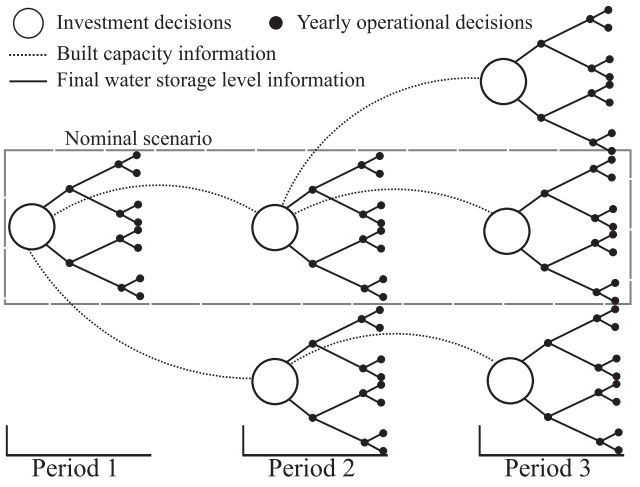

Fig. 1. Scenario structure with one *nominal* long-term scenario and two *extreme* long-term scenarios.

The proposed model is first described for only one long-term scenario in Section 2.1, where the notation and mathematical formulation are outlined. The model is then extended to accommodate multiple long-term scenarios in Section 2.2 through a compact notation.

2.1. Optimization model under one long-term scenario

Under each long-term scenario, the proposed optimization model is described by (1)–(14).

2.1.1. Objective function

$$\min \sum_{p\in\mathscr{P}} f_p \left\{ \left[\sum_{g\in\mathscr{G}} C^G_{g,p}(\phi^{Gfix}_{g,p} + \phi^{Ginv}_{g,p}) + \sum_{\ell\in\mathscr{L}} C^L_{\ell,p}(\phi^{Lfix}_{\ell,p} + \phi^{Linv}_{\ell,p}) \right] y_p + \sum_{h\in\mathscr{H}_p} \sum_{s\in\mathscr{S}} \theta_h \pi_s \left[\sum_{g\in\mathscr{G}} P_{g,h,s}(\phi^{OM}_g + \phi^{fuel}_{g,p}) \right] \right\} \tag{1}$$

The objective calculated in (1) is to minimize total investment and expected operational costs over all inflow scenarios. Annualized investment and fixed O&M costs are considered for generation and transmission capacity, which are multiplied by the number of years in each period. Variable O&M and fuel costs are calculated for power generation in each representative hour and inflow scenario, and are scaled up to the period and multiplied by its probability to calculate the expected costs. A discount factor is used to bring costs to present value and the amounts are summed up for all periods.

2.1.2. Power system operational constraints

$$l_{b,h} + \sum_{\ell\in\mathscr{L}^{out}_b} F_{\ell,h,s} + E_{b,h,s} = \sum_{g\in\mathscr{G}_b} P_{g,h,s} + \sum_{\ell\in\mathscr{L}^{in}_b} \eta^L_\ell F_{\ell,h,s} \quad \forall\, b\in\mathscr{B}, h\in\mathscr{H}, s\in\mathscr{S} \tag{2}$$

The power balance constraint is expressed in (2). Load at each bus and hour must be satisfied by local generators and power imports from other buses for every inflow scenario. Excess renewable power is allowed to be curtailed and no load shedding is permitted. Transmission is represented by a transport model. Losses are represented by a

constant loss factor that penalizes bus imports.

$$0 \leqslant F_{\ell,h,s} \leqslant C^L_{\ell,p} \quad \forall\, \ell \in \mathscr{L},\, h \in \mathscr{H}_p,\, s \in \mathscr{S},\, p \in \mathscr{P} \tag{3}$$

$$0 \leqslant P_{g,h,s} \leqslant C^G_{g,p}\, c_{g,h} \quad \forall\, g \in \mathscr{G},\, h \in \mathscr{H}_p,\, s \in \mathscr{S},\, p \in \mathscr{P} \tag{4}$$

Eqs. (3) and (4) constrain transmission flow and generator dispatch levels according to installed capacity. For variable generators, $c_{g,h}$ represents the capacity factor resulting from the amount of renewable resource present in each hour. Only one deterministic profile of renewable generation is considered for every year. For other generators, this parameter represents a derating factor to account for average availability.

$$P_{g,h+1,s} - P_{g,h,s} \leqslant r^{up}_g\, C^G_{g,p} \quad \forall\, g \in \mathscr{G},\, h \in \mathscr{H}_d,\, d \in \mathscr{D},\, s \in \mathscr{S} \tag{5a}$$

$$P_{g,h,s} - P_{g,h+1,s} \leqslant r^{dn}_g\, C^G_{g,p} \quad \forall\, g \in \mathscr{G},\, h \in \mathscr{H}_d,\, d \in \mathscr{D},\, s \in \mathscr{S} \tag{5b}$$

The use of representative days with hourly resolution allows the implementation of ramping constraints (5a) and (5b). Ramp rates are regarded as fractions of the current installed capacity for each project. The hours of each day are considered in a circular manner, so that dispatch in the last hour ramps to the first of the same day.

2.1.3. Power system construction constraints

$$C^G_{g,p} = b^G_{g,i} + \sum_{i=1}^{p-1} B^G_{g,i} \quad \forall\, g \in \mathscr{G},\, p \in \mathscr{P} - \{1\} \tag{6}$$

$$C^L_{\ell,p} = b^L_{\ell,i} + \sum_{i=1}^{p-1} B^L_{\ell,i} \quad \forall\, t \in \mathscr{L},\, p \in \mathscr{P} - \{1\} \tag{7}$$

Eqs. (6) and (7) show how cumulative installed capacity considers existing infrastructure plus new additions. Expansions that are decided on one period are considered to be available for operations from the next period onwards.

$$0 \leqslant B^G_{g,p} \leqslant \overline{B^G}_g \quad \forall\, g \in \mathscr{G},\, p \in \mathscr{P} \tag{8a}$$

$$0 \leqslant C^G_{g,p} \leqslant \overline{C^G}_g \quad \forall\, g \in \mathscr{G},\, p \in \mathscr{P} \tag{8b}$$

$$0 \leqslant B^L_{\ell,p} \leqslant \overline{B^L}_\ell \quad \forall\, \ell \in \mathscr{L},\, p \in \mathscr{P} \tag{9a}$$

$$0 \leqslant C^L_{\ell,p} \leqslant \overline{C^L}_\ell \quad \forall\, \ell \in \mathscr{L},\, p \in \mathscr{P} \tag{9b}$$

Investment in each period is limited in (8a) and (9a) to reflect labor, resource or capital mobilization bounds. Additionally, cumulative capacities are capped in (8b) and (9b) to reflect resource and terrain availability, or other factors.

2.1.4. Hydraulic system constraints

$$w_{n,h,s}\, \theta_h + V_{n,h,s} + \sum_{c \in \mathscr{C}^{in}_n} W_{c,h,s}\, \theta_h = V_{n,h+1,s} + \sum_{c \in \mathscr{C}^{out}_n} W_{c,h,s}\, \theta_h$$
$$\forall\, n \in \mathscr{N},\, h \in \mathscr{H},\, s \in \mathscr{S} \tag{10}$$

The water network is composed of *nodes*, which receive natural inflows and can store water, and *connections*, which transport water downstream between them. Conservation of mass at each node, hour and inflow scenario is enforced by (10). Water flows in each representative day are scaled up to the number of days represented.

$$\underline{V}_{n,h,s} \leqslant V_{n,h,s} \leqslant \overline{V}_{n,h,s} \quad \forall\, n \in \mathscr{N},\, h \in \mathscr{H},\, s \in \mathscr{S} \tag{11}$$

Water volume storage at each node is constrained by (11) to reflect design sizing, terrain limits, and regulatory requirements that vary between locations and seasons. Non-reservoir nodes have no storage capacity and act merely as junctions.

$$V^i_n = V_{n,first(\mathscr{H}_p),s} \quad \forall\, n \in \mathscr{N}^R,\, p \in \mathscr{P},\, s \in \mathscr{S} \tag{12a}$$

$$V^i_n \leqslant V_{n,last(\mathscr{H}_p),s} \quad \forall\, n \in \mathscr{N}^R,\, p \in \mathscr{P},\, s \in \mathscr{S} \tag{12b}$$

Using water from reservoirs in hydroelectric plants implies no immediate cost for the system, so a boundary condition must be set in order to prevent their excessive depletion. Previous works on hydrothermal EP force reservoirs to end each year with at least the same volume as in the beginning (e.g. see [18,30,4,15,9]). Constraints (12a) and (12b) extend that capacity to a multi-year horizon. The water volume of reservoirs at the first hour of each period —*first* ($\mathscr{H}_p$)—is given, and in the last hour of each period—*last* ($\mathscr{H}_p$)—the volume must reach at least the initial value.

$$P_{g,h,s}/\eta^H_g \leqslant W_{c,h,s} \quad \forall\, g \in \mathscr{G}^H,\, h \in \mathscr{H},\, s \in \mathscr{S} \tag{13}$$

Eq. (13) links the power system and the water network at hydroelectric generators. A constant hydraulic efficiency is assumed, in order to avoid non-linearities in the problem. Spilled water is cascaded downstream and can be used by other units.

$$V_{n,h,s} = V_{n,h,s'} \quad \forall\, n \in \mathscr{N},\, s' \in \Gamma_{h,s},\, h \in \mathscr{H},\, s \in \mathscr{S} \tag{14}$$

Dispatch decisions in hydrothermal systems must account for all possible scenarios that could follow, which is termed *non-anticipativity*. For example, a system operator would consider storing water in a year with high inflow availability in order to hedge the risk of future low inflow availability. Eq. (14) enforces such non-anticipativity by making water storage decisions in every hour to be the same in all inflow scenarios that are indistinguishable up to that moment. Since operational uncertainty is represented by water inflow scenarios and all other parameters are considered deterministic, constraining water levels is enough to ensure non-anticipativity.

2.2. Extension to multiple long-term scenarios

The linear program outlined in (1)–(14)can be written in the following compact manner:

$$\min_{\mathbf{x}} \quad \sum_{p \in \mathscr{P}} \left[c_p^\top \mathbf{x}_p + \sum_{s \in \mathscr{S}} \pi_s Q(p, \mathbf{x}_p, s) \right] \tag{15a}$$

$$\text{s. t.} \quad D\mathbf{x} \geqslant e \tag{15b}$$

$$\text{where} \quad Q(p, \mathbf{x}_p, s) = \min_{\mathbf{y}_{p,s}} \quad d_p^\top \mathbf{y}_{p,s} \tag{16a}$$

$$\text{s. t.} \quad W\mathbf{y}_{p,s} \geqslant u_p - T\mathbf{x}_p \tag{16b}$$

Here, (15a) summarizes (1) by representing capacity decisions—variables C^L, C^G, B^L, B^G—in period p by the vector $\mathbf{x}_p$, their investment costs by c_p and their constraints (6)–(9b) by (15b). Operational costs in each inflow scenario and period are represented by the function Q, which reflects the total cost of dispatch decisions —P, F, D, W, V—, represented by the vector $\mathbf{y}_{p,s}$. In turn, (16b) condenses constraints (2)–(5b) and (10)–(14). The problem can then be generalized to accommodate multiple long-term scenarios as follows.

$$\min_{\mathbf{x}} \quad \sum_{\omega \in \Omega} \gamma_\omega \left\{ \sum_{p \in \mathscr{P}} \left[c_p^{\omega\top} \mathbf{x}_p^\omega + \sum_{s \in \mathscr{S}} \pi_s Q(p, \mathbf{x}_p^\omega, s, \omega) \right] \right\} \tag{17a}$$

$$\text{s. t.} \quad (\mathbf{x}_p^\omega, \mathbf{y}_{p,s}^\omega) \in \Delta_p^\omega \tag{17b}$$

$$\mathbf{x}_p^\omega = \mathbf{x}_p^{\omega'}\ \forall\, \omega' \in \Psi_p^\omega,\, p \in \mathscr{P},\, \omega \in \Omega \tag{17c}$$

Let the set of all long-term scenarios be Ω, indexed by ω, and their probabilities be γ_ω. The objective function of the multiple scenario extension can then be expressed as in (17a). If Δ_p is the space of all combinations of $\mathbf{x}_p$ and $\mathbf{y}_{p,s}$ that satisfy constraints (15b) and (16b) in a given period p, then Δ_p^ω is the space of feasible combinations of $\mathbf{x}_p^\omega$ and $\mathbf{y}_{p,s}^\omega$ for long-term scenario ω, enforced in (17b). Different parameters and cost coefficients may be specified in each long-term scenario. Non-

anticipativity in investments is enforced in (17c), where construction decisions in each period are forced to be equal for all long-term scenarios that are indistinguishable up to that moment, gathered in the set Ψ_p^ω.

As discussed in Section 1, multiple works on EP formulate long-term scenario trees and assign probabilities according to different methods so that obtained investment plans have the least expected cost. However, assigning discrete probabilities to scenarios that span several decades, when unfolding phenomena are not completely understood, may become an insurmountable challenge.

This work captures long-term uncertainty by minimizing investment and operations costs for the most likely to happen scenario, termed *nominal* long-term scenario and highlighted in Fig. 1. In addition, *extreme* long-term scenarios that represent high-impact phenomena are considered. The probabilities of the latter are considered to be near-zero, so the optimization problem (17a)–(17c)ignores their costs. Hence, these extreme scenarios only add constraints that prepare the system to meet the required load in a reliable manner in such cases. However, only considering feasibility in those extreme long-term scenarios could result in investment plans which imply unacceptably high operational costs. A cap on total costs—ϕ_ω^{cap}—is then implemented for each extreme long-term in the additional constraint (17d), so that the decision maker can set the maximum cost it is willing to accept.

$$\sum_{p\in\mathscr{P}}\left[c_p^{\omega\top}\mathbf{x}_p^\omega+\sum_{s\in\mathscr{S}}\pi_s Q(p,\mathbf{x}_p^\omega,s,\omega)\right]\leqslant\phi_\omega^{\text{cap}}\quad\forall\,\omega\in\Omega \tag{17d}$$

3. Solution method

The complete problem structure defined in (17a)–(17d)consists on several long-term scenarios linked by investment decisions. The problem is decomposed and solved using the Progressive Hedging Algorithm (PHA) proposed by Rockafellar and Wets [29]. This method has received significant interest recently, because it can take advantage of modern parallel and distributed computing to reduce solution times for large problems. It has been successfully applied to power system problems, such as for stochastic EP with uncertainty of load and renewable output profiles by Muñoz and Watson [20], and for reservoir management in operations coordination by Dos Santos et al. [5].

Algorithm 1 gives an overview on the PHA, where $f(\mathbf{x})$ represents the objective function in (15a). It is an augmented Lagrangian method that decomposes the problem on a long-term scenario basis by relaxing the non-anticipativity constraint (17c). The objective function of each long-term scenario subproblem is then augmented with penalizing factors that steer each solution toward a non-anticipative optimum.

Algorithm 1. Progressive Hedging.

1: $i\leftarrow 0$, gap $\leftarrow\infty$
2: $\mathbf{w}_i^\omega\leftarrow\mathbf{0}\ \ \forall\,\omega\in\Omega$
3: $\mathbf{x}_i^\omega\leftarrow\text{argmin}_{\mathbf{x}}[f(\mathbf{x}_i^\omega)]$ s.t. [(15b)–(16b)]
4: **while** gap $\geqslant\epsilon$ **do**
5: $\quad\overline{\mathbf{x}}_i\leftarrow\sum_{\omega\in\Omega}\gamma^\omega\mathbf{x}_i^\omega$
6: $\quad i\leftarrow i+1$
7: $\quad\mathbf{w}_i^\omega\leftarrow\mathbf{w}_{i-1}^\omega+\boldsymbol{\rho}^\top(\mathbf{x}_{i-1}^\omega-\overline{\mathbf{x}}_{i-1})\ \ \forall\,\omega\in\Omega$
8: $\quad\mathbf{x}_i^\omega\leftarrow\text{argmin}_x[f(\mathbf{x}_i^\omega)+\mathbf{w}_i^{\omega\top}\mathbf{x}_i^\omega+\frac{1}{2}\boldsymbol{\rho}^\top\|\mathbf{x}_i^\omega-\overline{\mathbf{x}}_i)\|^2]$ s.t. [(15b)–(16b)]
9: $\quad\text{gap}\leftarrow\sum_{\omega\in\Omega}\|\mathbf{x}_i^\omega-\overline{\mathbf{x}}_i\|^2$

The proposed model includes extreme long-term scenarios with near-zero probabilities and a nominal long-term scenario with a near-one probability. This raises a problem for the application of the PHA. Having scenarios with widely different probabilities increases solving time significantly, because the weighted averaged value of each variable, calculated as in Line 5 in each iteration, will result in a value excessively close to that of the scenario with high probability. In consequence, penalization factors for the nominal long-term scenario calculated on Line 7 will be light. Thus, the solution of that scenario will only slightly vary in each iteration, increasing the number of iterations needed for convergence.

An alternative formulation drawn from [2] is implemented, which is also proven to achieve convergence and optimality for convex problems. Instead of calculating a probability-weighted average, Line 5 can be replaced by a simple arithmetic average as in (18). The penalty factor ρ for all variables must be multiplied by the inverse of the probability of their scenario, so convergence to the optimum is maintained. Line 7 in the original algorithm is then replaced by (19).

$$\overline{\mathbf{x}}_i\leftarrow\frac{\sum_{\omega\in\Omega}\mathbf{x}_i^\omega}{|\Omega|} \tag{18}$$

$$\mathbf{w}_i^\omega\leftarrow\mathbf{w}_{i-1}^\omega+\frac{1}{\gamma^\omega}\boldsymbol{\rho}^\top(\mathbf{x}_{i-1}^\omega-\overline{\mathbf{x}}_{i-1})\ \ \forall\,\omega\in\Omega \tag{19}$$

4. Implementation and computational experiments

The proposed model was implemented by developing new modules in the open source SWITCH platform [7], which is a planning tool based in Python/Pyomo and publicly available.[1] Linear subproblems are solved using Gurobi 7.0 with a 0.01% duality gap on an Intel Xeon E5-2620 24-core machine with 32 GB of memory.

All case studies comprise ten two-year investment periods spanning from 2020 to 2039 on the future interconnected power system of Chile. As of 2016, it consisted of two separate systems serving an aggregate demand of 71.68 TWh/year with nearly 20 GW of installed capacity. Future load was projected using estimated load growth rates. Renewable generation profiles were based on historical data. Existing and proposed generators and transmission lines were reduced to a total of 68 aggregated generators and 23 lines that connect 20 buses. The technologies associated with each generation project are described in Table 1. Three types of hydroelectric plants are modeled: those located at a dam (Reservoir), those located downstream from a dam (Series), and run-of-river (RoR).

In order to represent inflow uncertainty, one inflow scenario tree was used in all two-year periods, which was constructed from historical yearly records. The 56 current historical records (1960–2015) were clustered into 3 representative years using hierarchical clustering and Dynamic Time Warping as a distance metric [18]. Aravena and Gil [1] show that hydrological timeseries in Chile present negligible annual autocorrelation; thus, each year's hydrology can be assumed independent of previous years. In consequence, the 3 representative years can be combined into 9 different trajectories for each two-year period. An additional trajectory with null probability is added to each scenario tree to enforce feasibility under a drought scenario, so that a capacity margin is considered in the EP process.

Representative days were chosen by clustering hourly net load (total load – wind and solar generation) curves of every day of 2015. The 10% of days with the most abrupt evening ramps were clustered together and the remaining 90% were clustered into 3 groups by using hierarchical clustering with the Dynamic Time Warping distance metric. The representative day from each cluster was chosen based on a minimum variance criterion. The use of 4 representative days yields a total of 96 h per year.

Sections 4.1 and 4.2 exhibit two case studies that consider only one

[1] The on-line public repository where SWITCH is published can be found at: https://github.com/switch-model/switch.

Table 1
Characteristics of generator technologies.

Technology	Plants	Investment Cost [US$/kW]	Average Fuel Cost in 2020 [US$/MWh]	Average Hourly Ramp Rate [fraction of capacity]
RoR	8	3100	–	1.0
Series	5	3400	–	1.0
Reservoir	12	3100	–	1.0
Wind	11	2100	–	1.0
Solar PV	9	1950	–	1.0
Diesel Gas Turbine	6	946	200	0.5
Combined Cycle Gas Turbine	5	1100	89	0.3
Diesel Motor Facility	5	910	166	1.0
Coal Steam Turbine	7	3000	41	0.0

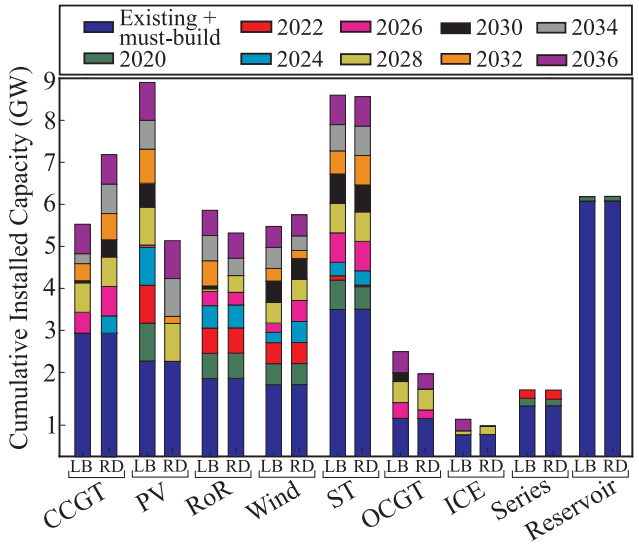

Fig. 2. Generation investment plans yielded by the LB and RD models for 2.5 GW of must-build reservoir hydro.

Table 2
Total Cost and Operational metrics of LB and RD investment plans.

Must-build Hydro (GW)	Total Cost (Billion 2015US$)			Load Shed (%)		Curtailed Energy (%)	
	LB	RD	Δ%	LB	RD	LB	RD
0	35.63	34.98	1.88	0.72	0.02	1.96	0.12
2.5	32.25	31.20	3.38	0.73	0.01	1.45	0.20
3.75	30.75	29.86	2.98	0.68	0.00	1.32	0.32

Table 3
Total investment costs (Billion 2015US$).

Discount rate	NA	IA		PF	
0%	37.63	38.27	+1.69%	37.61	−0.07%
2%	46.25	46.92	+1.46%	46.24	−0.01%
4%	37.67	38.25	+1.56%	37.69	+0.06%
7%	28.06	28.50	+1.56%	28.05	−0.04%
10%	21.07	21.47	+1.87%	21.07	−0.02%
12%	17.56	17.91	+2.00%	17.55	−0.01%
14%	14.70	14.99	+2.03%	14.69	−0.03%

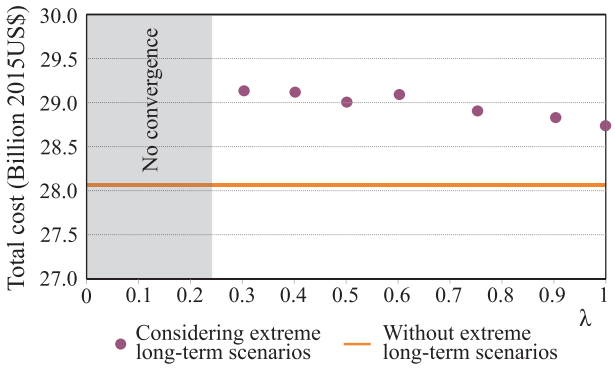

Fig. 3. Total cost in the nominal long-term scenario with and without considering extreme scenarios for multiple cost caps.

long-term scenario using the model in (1)–(14). Long-term uncertainty is considered in Section 4.3.

4.1. Case study 1—Representative days versus load blocks

The effect of time resolution is examined by comparing the outputs of a model that considers representative days with hourly resolution—the *RD model*—with those of an equivalent model that considers monthly load blocks—the *LB model*. The RD model is completely described by (1)–(14). The LB model is analogous, but hours are replaced by load blocks, ramping constraints (5a)–(5b) are ignored, and water storage balancing is performed on a monthly basis instead of hourly in (10). Inflows, renewable resource availability and load profiles are discretized into 8 blocks per month in the LB model, for a total of 96 blocks per year.

In order to compare the LB and RD models for different levels of flexibility in the system, each model is run with different portfolios of new must-build reservoir hydro power plants equivalent to 3.75 GW, 2.5 GW, and 0 GW of additional capacity. Investment plans for the 2.5 GW case can be observed in Fig. 2. The wall-clock run times in this case were 627 s for the LB model and 1102 s for the RD model. It can be observed that the LB model delivers a construction plan with more investment on PV and RoR projects than the RD model, which in turn builds more combined cycle units.

The operational performance of the investment plans obtained by the LB and RD models are assessed using an economic dispatch model with ramping constraints, spanning all the hours of the 2022–2039 horizon and all inflow scenarios in each period. To avoid infeasibility, load shedding is allowed at a cost of 500 US$/MWh. Cost differences and operational metrics related to efficiency and reliability are reported in Table 2 for these simulations.

The investment plan produced by the RD model yields between 1.88% and 3.38% smaller total costs (investment and operation) than the LB plan depending on the amount of must-build hydro power. Moreover, curtailed energy throughout the operational horizon can reach up to 1.96% of total generation with the LB plan, whereas it is 4–17 times smaller in the RD plan. Similarly, load shed when operating the RD investment plan is practically null, whereas the LB plan contemplates shedding around 0.7% of the load in all cases. These results show that representative days offer a better representation of operational conditions than load blocks, given their ability to capture intra-day flexibility requirements. This becomes clearer in cases with less reservoir hydro capacity—and therefore less flexible resources—in which the differences between the two approaches are more apparent.

4.2. Case study 2—The role of inter-annual water storage

As discussed in Section 2, most existing hydrothermal EP models do not allow inter-annual water storage in reservoirs and, thus,

underestimate the ability of the system to hedge against inflow uncertainty by storing water throughout successive years. To explore the impact of allowing this feature in a capacity planning model under uncertainty, three variations of the model described in Section 2.1 are studied. The first formulation (IA) only allows *intra-annual* storage, as existing models in literature. The second formulation, called *perfect foresight* (PF), allows inter-annual storage, but ignores the non-anticipativity constraint (14). Finally, the complete model (NA) allows inter-annual storage and includes the *non-anticipativity* constraints. In order to also study the timing of investment decisions, these problems are solved using multiple discount rates (reflected in parameter f_p). Total investment costs are presented in Table 3.

The IA model incurs in 1.56% higher investment costs than the NA formulation when the discount rate is set at its nominal value of 7%, but as the value of this parameter increases, cost differences increase to up to 2.03%. Results show that this is caused by the timing of constructions, since capacity addition decisions are similar, but slightly shifted in time. The IA model yields a plan with earlier investments given the need for more capacity in dry years, whereas the other models can compensate this need by storing water during years of high inflows. However, capturing the inflow uncertainty hedging capability of reservoirs causes model run time to increase from 120 s for the IA model, to 297 s for the PF model due to constraints linking different scenarios throughout a period. Additionally, results indicate that enforcing non-anticipativity under this dataset does not produce significantly different investment plans, since cost difference did not surpass 0.1% for any discount rate.

4.3. Case study 3—Robustness of the expansion plan

The PHA described in Section 3 is implemented by modifying the PySP Pyomo module. The values of ρ are set according to the *cost proportional* strategy proposed by Watson and Woodruff [32]. Near-zero probabilities for extremes scenarios are set to $\gamma = 10^{-4}$ and penalty factors (ρ) associated with their variables are amplified by a factor of 10^4.

Five extreme long-term scenarios are considered to represent high-impact low-probability phenomena. These scenarios include a 25% increase in fuel and investment costs of combustion technologies, as well as a 50% reduction in all water inflows. The PHA is set to stop at a maximum of 12 iterations or when the convergence metric $\mathbf{x}^c = \sum_{\omega\in\Omega}\sum_{i\in I}|\mathbf{x}_i^{\omega} - \bar{\mathbf{x}}_i|$ satisfies $\mathbf{x}^c \leqslant 1200$ MW, where 1200 MW represents less than 1% of the total capacity installed in all long-term scenarios.

Eq. (17d) in Section 2.2 establishes caps for the total costs under each extreme long-term scenario. A minimum and a maximum cost cap ($\phi_\omega^{\min}$ and $\phi_\omega^{\max}$) are determined for each long-term scenario. The minimum caps $\phi_\omega^{\min}$ are obtained by calculating the optimal planning and operations for each scenario ω deterministically. The maximum caps $\phi_\omega^{\max}$ are obtained by calculating the total operational cost in each scenario ω using the investment plan obtained in the nominal long-term scenario. Multiple caps are experimented with, calculated as a linear combinations of their minimum and maximum values through different values of λ, as follows: $\phi_\omega^{\text{cap}} = \lambda\phi_\omega^{\min} + (1-\lambda)\phi_\omega^{\max}$. Fig. 3 reports the total cost incurred in the nominal long-term scenario for various values of λ.

It can first be noted that the PHA provides a solution with 2.5% higher cost than the deterministic form when the value of λ is 1—which is the equivalent of ignoring the extreme long-term scenarios. This reflects the optimality gap remaining when the iteration limit or the convergence criterion are met in PHA.

A key insight is that most of the additional costs associated with extreme long-term scenarios can be avoided with marginal additional investments. In particular, Fig. 3 shows that by accepting an approximately 1.2% cost increase under the nominal long-term scenario, the system can reduce 70% of the additional costs incurred in all the extreme long-term scenarios.

5. Conclusions

This work presents an expansion planning model that can handle large scale integration of VRE sources in hydrothermal systems. Operations are represented with an hourly resolution using selected representative days, the water network is explicitly modeled, and yearly inflow uncertainty is endogenously accounted for. In addition, uncertainty in the investment scale is considered through extreme scenarios.

Numerical experiments on the Chilean power system show that the use of representative days for hourly operations yield a more flexible plan than with the use of load blocks. The system performs better both in terms of economic and reliability metrics, such as load shed and curtailed energy. In addition, the proposed model highlights the role of reservoir hydro-power in providing intra-day flexibility.

The results also illustrate the capability of reservoirs for hedging the system against inflow uncertainty by storing water throughout successive years. In the long-term, computational experiments indicate that small changes in the nominal investment plan might help mitigate most of the risk associated with extreme scenarios. The PHA is successfully applied with near-zero probability scenarios for an actual system, allowing a more efficient use of available computational resources.

Acknowledgments

The authors thank Josiah Johnston for his valuable comments on this research and the writing of this paper.

This research was partially funded by the Complex Engineering Systems Institute, ISCI (CONICYT: FB0816), CORFO/13CEI2-21803, CONICYT/FONDECYT/11170630 and by SERC–Chile (CONICYT/FONDAP/15110019).

B. Maluenda also acknowledges the financial support of the National Commission for Scientific and Technological Research under Scholarship CONICYT-PCHA/MagísterNacional/2016–22161429.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijepes.2018.06.008.

References

[1] Aravena I, Gil E. Hydrological scenario reduction for stochastic optimization in hydrothermal power systems. Appl Stochastic Models Bus Ind 2015;31(2):231–40.
[2] Birge JR, Louveaux F. Introduction to stochastic programming. second ed. Springer series in operations research New York, NY, USA: Springer; 2011.
[3] Campodónico N, Binato S, Pereira M, Mayak F. Expansion planning of generation and interconnections under uncertainty. 3rd Balkans Power Conf; 2003.
[4] Costa J, Campodónico N, Gorenstin B, Pereira M. A model for optimal energy expansion in interconnected hydrosystems. In: Proc 10th Power Syst Comput Conf; 1990. p. 40–7.
[5] Dos Santos ML, da Silva EL, Finardi EC, Goncalves RE. Practical aspects in solving the medium-term operation planning problem of hydrothermal power systems by using the progressive hedging method. Int J Elect Power Energy Syst 2009;31(9):546–52.
[6] Feng Y, Ryan SM. Scenario construction and reduction applied to stochastic power generation expansion planning. Comput Oper Res 2013;40(1):9–23.
[7] Fripp M. Switch: a planning tool for power systems with large shares of intermittent renewable energy. Environ Sci Technol; 2012.
[8] Gabriel SA, Conejo AJ, Fuller JD, Hobbs BF, Ruiz C. Complementarity modeling in energy markets. New York: Springer-Verlag; 2013.
[9] Gil E, Aravena I, Cárdenas R. Generation capacity expansion planning under hydro uncertainty using stochastic mixed integer programming and scenario reduction. IEEE Trans Power Syst 2015;30(4):1838–47.
[10] Hemmati R, Hooshmand RA, Khodabakhshian A. Comprehensive review of generation and transmission expansion planning. IET Gener Transm Distrib 2013;7(9):955–64.
[11] Hobbs BF, Xu Q, Ho J, Donohoo P, Kasina S, Ouyang J, et al. Adaptive transmission planning: Implementing a new paradigm for managing economic risks in grid expansion. IEEE Power Energy Mag 2016;14(4):30–40.
[12] Jin S, Botterud A, Ryan SM. Temporal versus stochastic granularity in thermal generation capacity planning with wind power. IEEE Trans Power Syst 2014;29(5):2033–41.

[13] Jonghe CD, Hobbs BF, Belmans R. Optimal generation mix with short-term demand response and wind penetration. IEEE Trans Power Syst 2012;27(2):830–9.
[14] Kagiannas AG, Askounis DT, Psarras J. Power generation planning: a survey from monopoly to competition. Int J Elect Power Energy Syst 2004;26(6):413–21.
[15] Kenfack F, Guinet A, Ngundam JM. Investment planning for electricity generation expansion in a hydro dominated environment. Int J Energy Res 2001;25(10):927–37.
[16] Khodaei A, Shahidehpour M, Wu L, Li Z. Coordination of short-term operation constraints in multi-area expansion planning. IEEE Trans Power Syst 2012;27(4):2242–50.
[17] Li S, Coit DW, Felder F. Stochastic optimization for electric power generation expansion planning with discrete climate change scenarios. Electric Power Syst Res 2016;140:401–12.
[18] Liao TW. Clustering of time series data: a survey. Pattern Recogn 2005;38(11):1857–74.
[19] Massé P, Gibrat R. Application of linear programming to investments in the electric power industry. Manage Sci 1957;3(2):149–66.
[20] Muñoz FD, Watson J-P. A scalable solution framework for stochastic transmission and generation planning problems. Comput Manage Sci 2015;12(4):491–518.
[21] Munoz FD, Hobbs BF, Ho JL, Kasina S. An engineering-economic approach to transmission planning under market and regulatory uncertainties: Wecc case study. IEEE Trans Power Syst 2014;29(1):307–17.
[22] Nelson J, Johnston J, Mileva A, Fripp M, Hoffman I, Petros-Good A, Blanco C, Kammen DM. High-resolution modeling of the western north american power system demonstrates low-cost and low-carbon futures. Energy Policy 2012;43:436–47.
[23] Oliveira GC, Binato S, Pereira MVF. Value-based transmission expansion planning of hydrothermal systems under uncertainty. IEEE Trans Power Syst 2007;22(4):1429–35.
[24] Palmintier BS, Webster MD. Impact of operational flexibility on electricity generation planning with renewable and carbon targets. IEEE Trans Sustain Energy 2016;7(2):672–84.
[25] Park H, Baldick R. Stochastic generation capacity expansion planning reducing greenhouse gas emissions. IEEE Trans Power Syst 2015;30(2):1026–34.
[26] Pereira MVF, Pinto LMVG. Multi-stage stochastic optimization applied to energy planning. Math Programming 1991;52(1):359–75.
[27] Poncelet K, Delarue E, Six D, Duerinck J, D'haeseleer W. Impact of the level of temporal and operational detail in energy-system planning models. Appl Energy 2016;162:631–43.
[28] Poncelet K, Höschle H, Delarue E, Virag A, D'haeseleer W. Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems. IEEE Trans Power Syst 2017;32(3):1936–48.
[29] Rockafellar RT, Wets RJ-B. Scenarios and policy aggregation in optimization under uncertainty. Math Oper Res 1991;16(1):119–47.
[30] Sanghvi AP, Shavel IH. Investment planning for hydro-thermal power system expansion: stochastic programming employing the dantzig-wolfe decomposition principle. IEEE Trans Power Syst 1986;1(2):115–21.
[31] Vinasco G, Tejada D, Silva ED, Rider M. Transmission network expansion planning for the colombian electrical system: connecting the ituango hydroelectric power plant. Electric Power Syst Res 2014;110:94–103.
[32] Watson JP, Woodruff DL. Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. Comput Manage Sci 2011;8(4):355–70.
[33] Wogrin S, Galbally D, Reneses J. Optimizing storage operations in medium- and long-term power system models. IEEE Trans Power Syst 2016;31(4):3129–38.